

OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/935,144

DATE: 01/15/2002
 TIME: 17:40:23

Input Set : N:\Crif3\RULE60\09935144.raw
 Output Set: N:\CRF3\01152002\I935144.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Larsen, Glenn
 6 Sako, Dianne
 7 Chang, Xiao Jia
 8 Veldman, Geertruida M.
 9 Cumming, Dale
 10 Kumar, Ravindra
 11 Shaw, Gray

13 (ii) TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN

15 (iii) NUMBER OF SEQUENCES: 45

17 (iv) CORRESPONDENCE ADDRESS:

18 (A) ADDRESSEE: LEGAL AFFAIRS
 19 (B) STREET: 87 CAMBRIDGE PARK DRIVE
 20 (C) CITY: CAMBRIDGE
 21 (D) STATE: MA
 22 (E) COUNTRY: USA
 23 (F) ZIP: 02140

25 (v) COMPUTER READABLE FORM:

26 (A) MEDIUM TYPE: Floppy disk
 27 (B) COMPUTER: IBM PC compatible
 28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 29 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

31 (vi) CURRENT APPLICATION DATA:

C--> 32 (A) APPLICATION NUMBER: US/09/935,144
 C--> 33 (B) FILING DATE: 21-Aug-2001
 34 (C) CLASSIFICATION:

56 (vii) PRIOR APPLICATION DATA:

37 (A) APPLICATION NUMBER: 08/713,556
 38 (B) FILING DATE:
 41 (A) APPLICATION NUMBER: US 08/112,608
 42 (B) FILING DATE: 26-AUG-1993
 45 (A) APPLICATION NUMBER: PCT/US93/10168
 46 (B) FILING DATE: 22-OCT-1993
 49 (A) APPLICATION NUMBER: US 08/235,398
 50 (B) FILING DATE: 28-APR-1994
 53 (A) APPLICATION NUMBER: US 08/316,305
 54 (B) FILING DATE: 30-SEP-1994
 57 (A) APPLICATION NUMBER: US 08/428,734
 58 (B) FILING DATE: 25-APR-1995

60 (viii) ATTORNEY/AGENT INFORMATION:

61 (A) NAME: BROWN, SCOTT A.
 62 (B) REGISTRATION NUMBER: 32,724
 63 (C) REFERENCE/DOCKET NUMBER: GI 5213F

64 (ix) TELECOMMUNICATION INFORMATION:

65 (A) TELEPHONE: (617) 498-8224
 66 (B) TELEFAX: (617) 876-5851

RECEIVED
 FEB 06 2002
 TECH CENTER 1600/2900

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/935,144

DATE: 01/15/2002

TIME: 17:40:23

Input Set : N:\Crif3\RULE60\09935144.raw

Output Set: N:\CRF3\01152002\I935144.raw

69 (2) INFORMATION FOR SEQ ID NO: 1:

71 (i) SEQUENCE CHARACTERISTICS:

72 (A) LENGTH: 1649 base pairs

73 (B) TYPE: nucleic acid

74 (C) STRANDEDNESS: single

75 (D) TOPOLOGY: linear

77 (ii) MOLECULE TYPE: cDNA

79 (vi) ORIGINAL SOURCE:

80 (A) ORGANISM: Homo sapiens

81 (G) CELL TYPE: Promyelocyte

82 (H) CELL LINE: HL60

84 (vii) IMMEDIATE SOURCE:

85 (B) CLONE: PMT21:PL85

87 (ix) FEATURE:

88 (A) NAME/KEY: 5'UTR

89 (B) LOCATION: 1..59

91 (ix) FEATURE:

92 (A) NAME/KEY: CDS

93 (B) LOCATION: 60..1268

95 (ix) FEATURE:

96 (A) NAME/KEY: 3'UTR

97 (B) LOCATION: 1269..1649

100 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

102 GCCACTTCTT CTGGGCCCAC GAGGCAGCTG TCCCATGCTC TGCTGAGCAC GGTGGTGCC      59
104 ATG CCT CTG CAA CTC CTC CTG TTG CTG ATC CTA CTG GGC CCT GGC AAC      107
105 Met Pro Leu Gln Leu Leu Leu Leu Ile Leu Leu Gly Pro Gly Asn
106 1 5 10 15
108 AGC TTG CAG CTG TGG GAC ACC TGG GCA GAT GAA GCC GAG AAA GCC TTG      155
109 Ser Leu Gln Leu Trp Asp Thr Trp Ala Asp Glu Ala Glu Lys Ala Leu
110 20 25 30
112 GGT CCC CTG CTT GCC CGG GAC CGG AGA CAG GCC ACC GAA TAT GAG TAC      203
113 Gly Pro Leu Ala Arg Asp Arg Arg Gln Ala Thr Glu Tyr Glu Tyr
114 35 40 45
116 CTA GAT TAT GAT TTC CTG CCA GAA ACG GAG CCT CCA GAA ATG CTG AGG      251
117 Leu Asp Tyr Asp Phe Leu Pro Glu Thr Glu Pro Pro Glu Met Leu Arg
118 50 55 60
120 AAC AGC ACT GAC ACC ACT CCT CTG ACT GGG CCT GGA ACC CCT GAG TCT      299
121 Asn Ser Thr Asp Thr Thr Pro Leu Thr Gly Pro Gly Thr Pro Glu Ser
122 65 70 75 80
124 ACC ACT GTG GAG CCT GCT GCA AGG CGT TCT ACT GGC CTG GAT GCA GGA      347
125 Thr Thr Val Glu Pro Ala Ala Arg Arg Ser Thr Gly Leu Asp Ala Gly
126 85 90 95
128 GGG GCA GTC ACA GAG CTG ACC ACG GAG CTG GCC AAC ATG GGG AAC CTG      395
129 Gly Ala Val Thr Glu Leu Thr Thr Glu Leu Ala Asn Met Gly Asn Leu
130 100 105 110
132 TCC ACG GAT TCA GCA GCT ATG GAG ATA CAG ACC ACT CAA CCA GCA GCC      443
133 Ser Thr Asp Ser Ala Ala Met Glu Ile Gln Thr Thr Gln Pro Ala Ala
134 115 120 125
136 ACG GAG GCA CAG ACC ACT CCA CTG GCA GCC ACA GAG GCA CAG ACA ACT      491

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/935,144

DATE: 01/15/2002

TIME: 17:40:23

Input Set : N:\Crf3\RULE60\09935144.raw

Output Set: N:\CRF3\01152002\I935144.raw

137	Thr	Glu	Ala	Gln	Thr	Thr	Pro	Leu	Ala	Ala	Thr	Glu	Ala	Gln	Thr	Thr	
138		130					135					140					
140	CGA	CTG	ACG	GCC	ACG	GAG	GCA	CAG	ACC	ACT	CCA	CTG	GCA	GCC	ACA	GAG	539
141	Arg	Leu	Thr	Ala	Thr	Glu	Ala	Gln	Thr	Thr	Pro	Leu	Ala	Ala	Thr	Glu	
142	145						150					155				160	
144	GCA	CAG	ACC	ACT	CCA	CCA	GCA	GCC	ACG	GAA	GCA	CAG	ACC	ACT	CAA	CCC	587
145	Ala	Gln	Thr	Thr	Pro	Pro	Ala	Ala	Thr	Glu	Ala	Gln	Thr	Thr	Gln	Pro	
146							165					170				175	
148	ACA	GGC	CTG	GAG	GCA	CAG	ACC	ACT	GCA	CCA	GCA	GCC	ATG	GAG	GCA	CAG	635
149	Thr	Gly	Leu	Glu	Ala	Gln	Thr	Thr	Ala	Pro	Ala	Ala	Met	Glu	Ala	Gln	
150							180					185				190	
152	ACC	ACT	GCA	CCA	GCA	GCC	ATG	GAA	GCA	CAG	ACC	ACT	CCA	CCA	GCA	GCC	683
153	Thr	Thr	Ala	Pro	Ala	Ala	Met	Glu	Ala	Gln	Thr	Thr	Pro	Pro	Ala	Ala	
154							195					200				205	
156	ATG	GAG	GCA	CAG	ACC	ACT	CAA	ACC	ACA	GCC	ATG	GAG	GCA	CAG	ACC	ACT	731
157	Met	Glu	Ala	Gln	Thr	Thr	Gln	Thr	Thr	Ala	Met	Glu	Ala	Gln	Thr	Thr	
158		210					215					220					
160	GCA	CCA	GAA	GCC	ACG	GAG	GCA	CAG	ACC	ACT	CAA	CCC	ACA	GCC	ACG	GAG	779
161	Ala	Pro	Glu	Ala	Thr	Glu	Ala	Gln	Thr	Thr	Gln	Pro	Thr	Ala	Thr	Glu	
162	225						230					235				240	
164	GCA	CAG	ACC	ACT	CCA	CTG	GCA	GCC	ATG	GAG	GCC	CTG	CCC	ACA	GAA	CCC	827
165	Ala	Gln	Thr	Thr	Pro	Leu	Ala	Ala	Met	Glu	Ala	Leu	Ser	Thr	Glu	Pro	
166							245					250				255	
168	AGT	GCC	ACA	GAG	GCC	CTG	TCC	ATG	GAA	CCG	ACT	ACC	AAA	AGA	GGT	CTG	875
169	Ser	Ala	Thr	Glu	Ala	Leu	Ser	Met	Glu	Pro	Thr	Thr	Lys	Arg	Gly	Leu	
170							260					265				270	
172	TTC	ATA	CCC	TTT	TCT	GTG	TCC	TCT	GTT	ACT	CAC	AAG	GGC	ATT	CCC	ATG	923
173	Phe	Ile	Pro	Phe	Ser	Val	Ser	Ser	Val	Thr	His	Lys	Gly	Ile	Pro	Met	
174							275					280				285	
176	GCA	GCC	AGC	AAT	TTG	TCC	GTC	AAC	TAC	CCA	GTG	GGG	GCC	CCA	GAC	CAC	971
177	Ala	Ala	Ser	Asn	Leu	Ser	Val	Asn	Tyr	Pro	Val	Gly	Ala	Pro	Asp	His	
178		290					295					300					
180	ATC	TCT	GTG	AAG	CAG	TGC	CTG	CTG	GCC	ATC	CTA	ATC	TTG	GCG	CTG	GTG	1019
181	Ile	Ser	Val	Lys	Gln	Cys	Leu	Leu	Ala	Ile	Leu	Ile	Leu	Ala	Leu	Val	
182	305						310					315				320	
184	GCC	ACT	ATC	TTC	TTC	GTG	TGC	ACT	GTG	GTG	CTG	GCG	GTC	CGC	CTC	TCC	1067
185	Ala	Thr	Ile	Phe	Phe	Val	Cys	Thr	Val	Val	Leu	Ala	Val	Arg	Leu	Ser	
186							325					330				335	
188	CGC	AAG	GGC	CAC	ATG	TAC	CCC	GTG	CGT	AAT	TAC	TCC	CCC	ACC	GAG	ATG	1115
189	Arg	Lys	Gly	His	Met	Tyr	Pro	Val	Arg	Asn	Tyr	Ser	Pro	Thr	Glu	Met	
190							340					345				350	
192	GTC	TGC	ATC	TCA	TCC	CTG	TTG	CCT	GAT	GGG	GGT	GAG	GGG	CCC	TCT	GCC	1163
193	Val	Cys	Ile	Ser	Ser	Leu	Leu	Pro	Asp	Gly	Gly	Glu	Gly	Pro	Ser	Ala	
194							355					360				365	
196	ACA	GCC	AAT	GGG	GGC	CTG	TCC	AAG	GCC	AAG	AGC	CCG	GGC	CTG	ACG	CCA	1211
197	Thr	Ala	Asn	Gly	Gly	Leu	Ser	Lys	Ala	Lys	Ser	Pro	Gly	Leu	Thr	Pro	
198		370					375					380					
200	GAG	CCC	AGG	GAG	GAC	CGT	GAG	GGG	GAT	GAC	CTC	ACC	CTG	CAC	AGC	TTC	1259
201	Glu	Pro	Arg	Glu	Asp	Arg	Glu	Gly	Asp	Asp	Leu	Thr	Leu	His	Ser	Phe	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/935,144

DATE: 01/15/2002

TIME: 17:40:23

Input Set : N:\Crf3\RULE60\09935144.raw

Output Set: N:\CRF3\01152002\I935144.raw

```

202 385          390          395          400
204 CTC CCT TAGCTCACTC TGCCATCTGT TTTGGCAAGA CCCCACCTCC ACGGGCTCTC 1315
205 Leu Pro
208 CTGGGCCACC CCTGAGTGCC CAGACCCCCAA TCCACAGCTC TGGGCTTCTC CGGAGACCCC 1375
210 TGGGGATGGG GATCTTCAGG GAAGGAACTC TGGCCACCCA AACAGGACAA GAGCAGCCTG 1435
212 GGGCCAAGCA GACGGGCAAG TGGAGCCACC TCTTTCTCTC CTCCGGGGAT GAAGCCCAGC 1495
214 CACATTTCAG CCGAGGTCCA AGGCAGGAGG CCATTTACTT GAGACAGATT CTCTCCTTTT 1555
216 TCCTGTCCCC CATCTTCTCT GGGTCCCTCT AACATCTCCC ATGGCTCTCC CCGCTTCTCC 1615
218 TGGTCACTGG AGTCTCTCTC CCATGTACCC AAGG 1649
221 (2) INFORMATION FOR SEQ ID NO: 2:
223 (i) SEQUENCE CHARACTERISTICS:
224 (A) LENGTH: 402 amino acids
225 (B) TYPE: amino acid
226 (D) TOPOLOGY: linear
228 (ii) MOLECULE TYPE: protein
230 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
232 Met Pro Leu Gln Leu Leu Leu Leu Ile Leu Leu Gly Pro Gly Asn
233 1 5 10 15
235 Ser Leu Gln Leu Trp Asp Thr Trp Ala Asp Glu Ala Glu Lys Ala Leu
236 20 25 30
238 Gly Pro Leu Leu Ala Arg Asp Arg Arg Gln Ala Thr Glu Tyr Glu Tyr
239 35 40 45
241 Leu Asp Tyr Asp Phe Leu Pro Glu Thr Glu Pro Pro Glu Met Leu Arg
242 50 55 60
244 Asn Ser Thr Asp Thr Thr Pro Leu Thr Gly Pro Gly Thr Pro Glu Ser
245 65 70 75 80
247 Thr Thr Val Glu Pro Ala Ala Arg Arg Ser Thr Gly Leu Asp Ala Gly
248 85 90 95
250 Gly Ala Val Thr Glu Leu Thr Thr Glu Leu Ala Asn Met Gly Asn Leu
251 100 105 110
253 Ser Thr Asp Ser Ala Ala Met Glu Ile Gln Thr Thr Gln Pro Ala Ala
254 115 120 125
256 Thr Glu Ala Gln Thr Thr Pro Leu Ala Ala Thr Glu Ala Gln Thr Thr
257 130 135 140
259 Arg Leu Thr Ala Thr Glu Ala Gln Thr Thr Pro Leu Ala Ala Thr Glu
260 145 150 155 160
262 Ala Gln Thr Thr Pro Pro Ala Ala Thr Glu Ala Gln Thr Thr Gln Pro
263 165 170 175
265 Thr Gly Leu Glu Ala Gln Thr Thr Ala Pro Ala Ala Met Glu Ala Gln
266 180 185 190
268 Thr Thr Ala Pro Ala Ala Met Glu Ala Gln Thr Thr Pro Pro Ala Ala
269 195 200 205
271 Met Glu Ala Gln Thr Thr Gln Thr Thr Ala Met Glu Ala Gln Thr Thr
272 210 215 220
274 Ala Pro Glu Ala Thr Glu Ala Gln Thr Thr Gln Pro Thr Ala Thr Glu
275 225 230 235 240
277 Ala Gln Thr Thr Pro Leu Ala Ala Met Glu Ala Leu Ser Thr Glu Pro
278 245 250 255
280 Ser Ala Thr Glu Ala Leu Ser Met Glu Pro Thr Thr Lys Arg Gly Leu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/935,144

DATE: 01/15/2002

TIME: 17:40:23

Input Set : N:\Crif3\RULE60\09935144.raw

Output Set: N:\CRF3\01152002\I935144.raw

```

281          260          265          270
283 Phe Ile Pro Phe Ser Val Ser Ser Val Thr His Lys Gly Ile Pro Met
284          275          280          285
286 Ala Ala Ser Asn Leu Ser Val Asn Tyr Pro Val Gly Ala Pro Asp His
287          290          295          300
289 Ile Ser Val Lys Gln Cys Leu Leu Ala Ile Leu Ile Leu Ala Leu Val
290 305          310          315          320
292 Ala Thr Ile Phe Phe Val Cys Thr Val Val Leu Ala Val Arg Leu Ser
293          325          330          335
295 Arg Lys Gly His Met Tyr Pro Val Arg Asn Tyr Ser Pro Thr Glu Met
296          340          345          350
298 Val Cys Ile Ser Ser Leu Leu Pro Asp Gly Gly Glu Gly Pro Ser Ala
299          355          360          365
301 Thr Ala Asn Gly Gly Leu Ser Lys Ala Lys Ser Pro Gly Leu Thr Pro
302          370          375          380
304 Glu Pro Arg Glu Asp Arg Glu Gly Asp Asp Leu Thr Leu His Ser Phe
305 385          390          395          400
307 Leu Pro

```

311 (2) INFORMATION FOR SEQ ID NO: 3:

313 (i) SEQUENCE CHARACTERISTICS:

314 (A) LENGTH: 1239 base pairs

315 (B) TYPE: nucleic acid

316 (C) STRANDEDNESS: single

317 (D) TOPOLOGY: linear

319 (ii) MOLECULE TYPE: cDNA (synthetic)

321 (vi) ORIGINAL SOURCE:

322 (A) ORGANISM: Homo sapiens

323 (G) CELL TYPE: placenta

325 (ix) FEATURE:

326 (A) NAME/KEY: CDS

327 (B) LOCATION: 1..1239

330 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

332 ATG CCT CTG CAA CTC CTC CTG TTG CTG ATC CTA CTG GGC CCT GGC AAC      48
333 Met Pro Leu Gln Leu Leu Leu Leu Leu Ile Leu Leu Gly Pro Gly Asn
334 1          5          10          15
336 AGC TTG CAG CTG TGG GAC ACC TGG GCA GAT GAA GCC GAG AAA GCC TTG      96
337 Ser Leu Gln Leu Trp Asp Thr Trp Ala Asp Glu Ala Glu Lys Ala Leu
338          20          25          30
340 GGT CCC CTG CTT GCC CGG GAC CGG AGA CAG GCC ACC GAA TAT GAG TAC      144
341 Gly Pro Leu Leu Ala Arg Asp Arg Arg Gln Ala Thr Glu Tyr Glu Tyr
342          35          40          45
344 CTA GAT TAT GAT TTC CTG CCA GAA ACG GAG CCT CCA GAA ATG CTG AGG      192
345 Leu Asp Tyr Asp Phe Leu Pro Glu Thr Glu Pro Pro Glu Met Leu Arg
346 50          55          60
348 AAC AGC ACT GAC ACC ACT CCT CTG ACT GGG CCT GGA ACC CCT GAG TCT      240
349 Asn Ser Thr Asp Thr Thr Pro Leu Thr Gly Pro Gly Thr Pro Glu Ser
350 65          70          75          80
352 ACC ACT GTG GAG CCT GCT GCA AGG CGT TCT ACT GGC CTG GAT GCA GGA      288
353 Thr Thr Val Glu Pro Ala Ala Arg Arg Ser Thr Gly Leu Asp Ala Gly

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/935,144

DATE: 01/15/2002

TIME: 17:40:24

Input Set : N:\Crf3\RULE60\09935144.raw

Output Set: N:\CRF3\01152002\I935144.raw

L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]